

METHODS OF FABRICATING NITRIDE-BASED TRANSISTORS HAVING
REGROWN OHMIC CONTACT REGIONS AND NITRIDE-BASED
TRANSISTORS HAVING REGROWN OHMIC CONTACT REGIONS

FIELD OF THE INVENTION

The present invention relates to semiconductor devices and, more particularly, to transistors that incorporate nitride-based active layers.

5

BACKGROUND

The present invention relates to transistors formed of semiconductor materials that can make them suitable for high power, high temperature, and/or high frequency applications. Materials such as silicon (Si) and gallium arsenide (GaAs) have found
10 wide application in semiconductor devices for lower power and (in the case of Si) lower frequency applications. These, more familiar, semiconductor materials may not be well suited for higher power and/or high frequency applications, however, because of their relatively small bandgaps (*e.g.*, 1.12 eV for Si and 1.42 for GaAs at room temperature) and/or relatively small breakdown voltages.

15 In light of the difficulties presented by Si and GaAs, interest in high power, high temperature and/or high frequency applications and devices has turned to wide bandgap semiconductor materials such as silicon carbide (2.996 eV for alpha SiC at room temperature) and the Group III nitrides (*e.g.*, 3.36 eV for GaN at room temperature). These materials, typically, have higher electric field breakdown
20 strengths and higher electron saturation velocities as compared to gallium arsenide and silicon.

A device of particular interest for high power and/or high frequency applications is the High Electron Mobility Transistor (HEMT), which is also known as a modulation doped field effect transistor (MODFET). These devices may offer
25 operational advantages under a number of circumstances because a two-dimensional electron gas (2DEG) is formed at the heterojunction of two semiconductor materials with different bandgap energies, and where the smaller bandgap material has a higher electron affinity. The 2DEG is an accumulation layer in the undoped ("unintentionally doped"), smaller bandgap material and can contain a very high sheet electron

concentration in excess of, for example, 10^{13} carriers/cm². Additionally, electrons that originate in the wider-bandgap semiconductor transfer to the 2DEG, allowing a high electron mobility due to reduced ionized impurity scattering.

This combination of high carrier concentration and high carrier mobility can give the HEMT a very large transconductance and may provide a strong performance advantage over metal-semiconductor field effect transistors (MESFETs) for high-frequency applications.

High electron mobility transistors fabricated in the gallium nitride/aluminum gallium nitride (GaN/AlGaN) material system have the potential to generate large amounts of RF power because of the combination of material characteristics that includes the aforementioned high breakdown fields, their wide bandgaps, large conduction band offset, and/or high saturated electron drift velocity. A major portion of the electrons in the 2DEG is attributed to polarization in the AlGaN.

HEMTs in the GaN/AlGaN system have already been demonstrated. U.S. Patents 5,192,987 and 5,296,395 describe AlGaN/GaN HEMT structures and methods of manufacture. U.S. Patent No. 6,316,793, to Sheppard et al., which is commonly assigned and is incorporated herein by reference, describes a HEMT device having a semi-insulating silicon carbide substrate, an aluminum nitride buffer layer on the substrate, an insulating gallium nitride layer on the buffer layer, an aluminum gallium nitride barrier layer on the gallium nitride layer, and a passivation layer on the aluminum gallium nitride active structure.

One issue with the fabrication of nitride-based transistors involves the formation of ohmic contacts for such transistors. Conventionally, ohmic contacts have been formed through reactive ion etching (RIE) recesses for the contacts. However, without strict process control practices, RIE in nitride based materials may suffer from uniformity and reproducibility problems. Such problems could result in difficulty in controlling a fabrication process. Ohmic contacts that are formed without RIE have, typically, used high annealing temperatures (*e.g.* 900 °C). Such high annealing temperatures may damage the materials and/or the device.

SUMMARY OF THE INVENTION

Some embodiments of the present invention provide for fabrication of a transistor including forming a nitride-based channel layer on a substrate, forming a barrier layer on the nitride-based channel layer, forming a contact recess in the barrier

layer to expose a contact region of the nitride-based channel layer and forming a contact layer on the exposed contact region of the nitride-based channel layer using a low temperature deposition process. Fabrication may also include forming an ohmic contact on the contact layer and forming a gate contact disposed on the barrier layer adjacent the ohmic contact.

5 In further embodiments of the present invention, forming a contact layer on the exposed contact region of the nitride-based channel layer using a low temperature deposition process includes forming a contact layer by metal organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), plasma enhanced chemical
10 vapor deposition (PECVD), sputtering and/or hydride vapor phase epitaxy (HVPE). Furthermore, the low temperature deposition process may be a process other than mass transport from a wafer on which the transistor is formed.

In additional embodiments of the present invention, fabrication of the transistor further includes forming a first dielectric layer on the barrier layer and
15 forming a recess in the first dielectric layer. Forming a gate contact includes forming a gate contact in the recess. Forming a contact recess includes forming a contact recess in the first dielectric layer and the barrier layer that exposes a portion of the nitride-based channel layer. In other embodiments of the present invention, the gate contact may be formed on the first dielectric layer.

In still further embodiments of the present invention, the first dielectric layer comprises a silicon nitride layer. The silicon nitride layer may provide a passivation layer for the transistor.

In additional embodiments of the present invention, the contact recess extends into the channel layer. Furthermore, forming an ohmic contact may include forming an ohmic contact without annealing the ohmic contact. Forming an ohmic contact could include patterning a metal layer on the contact layer and annealing the patterned metal layer at a temperature of about 850 °C or less.

20 In other embodiments of the present invention, forming a contact layer on the exposed portions of the nitride-based channel layer includes forming a contact layer on the exposed portion of the nitride-based channel layer to a thickness sufficient to provide a sheet resistivity of less than a sheet resistivity of a two-dimensional electron gas region formed at an interface between the channel layer and the barrier layer.
25 Forming a contact layer may include forming n-type an InGaN, AlInN, AlInGaN and/or InN layer. In some embodiments of the present invention, the n-type nitride-

based layer formed is GaN and/or AlGaN. The InGaN, GaN, AlGaN, AlInN, AlInGaN and/or InN layer may be doped with Si, Ge and/or O during formation.

In some embodiments of the present invention, the contact layer includes an n-type degenerate semiconductor material other than GaN and AlGaN. The contact
5 layer may include a non-nitride Group III-V semiconductor material, a Group IV semiconductor material and/or a group II-VI semiconductor material.

In additional embodiments of the present invention, fabrication of the transistor further includes forming sidewalls of the channel layer to provide an increased surface area interface between the channel layer and the n-type contact layer
10 as compared to a planar interface. Forming an ohmic contact on the contact layer may include forming an ohmic contact on the contact layer that extends onto a portion of the channel layer or that terminates before the sidewall of the channel layer.

In further embodiments of the present invention, fabrication of the transistor includes forming holes in the channel layer adjacent the contact regions and placing
15 n-type nitride-based semiconductor material in the holes. Forming an ohmic contact on the contact layer further includes forming an ohmic contact on the contact layer and on the nitride-based semiconductor material in the holes.

In additional embodiments of the present invention, the contact layer extends onto the barrier layer.

20 In still other embodiments of the present invention, fabricating a transistor includes forming a nitride-based channel layer on a substrate, forming a barrier layer on the nitride-based channel layer, forming a masking layer on the barrier layer, patterning the masking layer and the barrier layer to provide contact opening that exposes a portion of the nitride-based channel layer, forming a contact layer on the
25 exposed portion of the nitride-based channel layer and the masking layer, selectively removing the masking layer and a portion of the contact layer on the masking layer to provide a contact region, forming an ohmic contact on the contact region and forming a gate contact disposed on the barrier layer adjacent the ohmic contact. Fabrication of the transistor may also include forming a first dielectric layer on the barrier layer and
30 forming a recess in the first dielectric layer. Forming a gate contact may include forming a gate contact in the recess. Forming a masking layer on the barrier layer may include forming a masking layer on the first dielectric layer. Patterning the masking layer and the barrier layer to provide contact openings that expose a portion of the nitride-based channel layer may include patterning the masking layer, the first

dielectric layer and the barrier layer to provide contact opening that exposes a portion of the nitride-based channel layer.

5 In certain embodiments of the present invention, the first dielectric layer includes a silicon nitride layer. The silicon nitride layer may provide a passivation layer for the transistor. The masking layer may be a dielectric layer. The dielectric layer may be a silicon oxide layer. The masking layer could be a photoresist masking layer.

10 Forming an ohmic contact may be provided by forming an ohmic contact without annealing the ohmic contact. Alternatively, forming an ohmic contact may be provided by patterning a metal layer on the contact region and annealing the patterned metal layer at a temperature of about 850 °C or less.

15 Forming a contact layer on the exposed portion of the nitride-based channel layer and the oxide layer may include forming a contact layer by metal organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), plasma enhanced chemical vapor deposition (PECVD), sputtering and/or hydride vapor phase epitaxy (HVPE). Forming a contact layer on the exposed portions of the nitride-based channel layer and the masking layer may be provided by forming a contact layer on the exposed portions of the nitride-based channel layer and the masking layer to a thickness sufficient to provide a sheet resistivity of less than a sheet resistivity of a 20 two-dimensional electron gas region formed at an interface between the channel layer and the barrier layer. Forming a contact layer may include forming an n-type InGa_N, AlInGa_N, InAlN and/or InN layer. In some embodiments, the nitride based contact layer may be GaN and/or AlGa_N. The InGa_N, AlInGa_N, InAlN, GaN, AlGa_N and/or InN layer may be doped with Si, Ge and/or O during formation.

25 In further embodiments of the present invention, fabrication of the transistor includes forming sidewalls of the channel layer to provide an increased surface area interface between the channel layer and the n-type contact layer compared to a planar interface. Forming an ohmic contact on the contact layer may include forming an ohmic contact on the contact layer that extends onto a portion of the channel layer or 30 that terminates before the sidewall of the channel layer. Additionally or alternatively, fabrication of the transistor may include forming holes in the channel layer adjacent the contact regions and placing a nitride-based semiconductor material in the holes. Forming an ohmic contact on the nitride-based contact region may include forming an

ohmic contact on the nitride-based contact region and on the nitride-based semiconductor material in the holes.

In other embodiments of the present invention, a high electron mobility transistor (HEMT) and methods of fabricating a HEMT are provided. The HEMT includes a nitride-based channel layer on a substrate, a barrier layer on the nitride-based channel layer, a contact recess in the barrier layer that extends into the channel layer, a contact region on the nitride-based channel layer in the contact recess, a gate contact disposed on the barrier layer. The contact region and the nitride-based channel layer include a surface area enlargement structure.

In some embodiments of the present invention, the surface area enlargement structure includes patterned sidewalls of portions of the contact recess that extends into the channel layer. In certain embodiments of the present invention, an ohmic contact is provided on the contact region that does not extend onto the channel layer in the area of the sidewalls. In other embodiments of the present invention, the ohmic contact extends onto the channel layer in the area of the sidewalls.

In additional embodiments of the present invention, the surface area enlargement structure includes holes extending into the channel layer with n-type nitride-based semiconductor material in the and the ohmic contact is in contact with the nitride-based semiconductor material in the holes.

The n-type nitride-based semiconductor material may include InN, AlGa_N, InGa_N, AlInGa_N, AlInN and/or GaN. The n-type nitride-based semiconductor material may be doped with Si, Ge and/or O. A silicon nitride layer may also be provided on the barrier layer and the gate contact may be provided in a recess in the silicon nitride layer.

Further embodiments of the present invention provide a high electron mobility transistor and methods of fabricating a transistor that includes a nitride-based channel layer on a substrate and a barrier layer on the nitride-based channel layer. At least one contact recess is provided in the barrier layer that extends into the channel layer. A region of metal and/or metal alloy is provided on the nitride-based channel layer in the contact recess to provide an ohmic contact. A gate contact disposed is on the barrier layer. The region of metal may extend onto the barrier layer.

Additional embodiments of the present invention provide a high electron mobility transistor and methods of fabricating a transistor that includes a nitride-based channel layer on a substrate and a barrier layer on the nitride-based channel layer. At

least one contact recess is provided in the barrier layer that extends into the channel layer. A region of n-type degenerate semiconductor material other than GaN or AlGaIn is provided on the nitride-based channel layer in the contact recess. An ohmic contact is provided on the region of n-type degenerate semiconductor material and a gate contact is disposed on the barrier layer. The region of n-type degenerate semiconductor material may extend onto the barrier layer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1G are schematic drawings illustrating fabrication of ohmic contacts in transistors according to embodiments of the present invention.

Figure 2 is a schematic illustration of transistors according to further embodiments of the present invention.

Figures 3A and 3B are schematic illustrations of transistors according to further embodiments of the present invention.

Figures 4A-4C are schematic illustrations of fabrication of transistors according to further embodiments of the present invention.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

The present invention now will be described more fully hereinafter with reference to the accompanying drawings, in which embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. In the drawings, the size and relative sizes of layers and regions may be exaggerated for clarity. It will be understood that when an element or layer is referred to as being "on", "connected to" or "coupled to" another element or layer, it can be directly on, connected or coupled to the other element or layer or intervening elements or layers may be present. In contrast, when an element is referred to as being "directly on", "directly connected to" or "directly coupled to" another element or layer, there are no intervening elements or layers present. Like numbers refer to like elements throughout. As used herein, the term "and/or" includes any and all combinations of one or more of the associated listed items.

It will be understood that, although the terms first, second, etc. may be used herein to describe various elements, components, regions, layers and/or sections, these

elements, components, regions, layers and/or sections should not be limited by these terms. These terms are only used to distinguish one element, component, region, layer or section from another region, layer or section. Thus, a first element, component, region, layer or section discussed below could be termed a second
5 element, component, region, layer or section without departing from the teachings of the present invention.

Furthermore, relative terms, such as "lower" or "bottom" and "upper" or "top," may be used herein to describe one element's relationship to another elements as illustrated in the Figures. It will be understood that relative terms are intended to
10 encompass different orientations of the device in addition to the orientation depicted in the Figures. For example, if the device in the Figures is turned over, elements described as being on the "lower" side of other elements would then be oriented on "upper" sides of the other elements. The exemplary term "lower", can therefore, encompasses both an orientation of "lower" and "upper," depending of the particular
15 orientation of the figure. Similarly, if the device in one of the figures is turned over, elements described as "below" or "beneath" other elements would then be oriented "above" the other elements. The exemplary terms "below" or "beneath" can, therefore, encompass both an orientation of above and below.

Embodiments of the present invention are described herein with reference to
20 cross-section illustrations that are schematic illustrations of idealized embodiments of the present invention. As such, variations from the shapes of the illustrations as a result, for example, of manufacturing techniques and/or tolerances, are to be expected. Thus, embodiments of the present invention should not be construed as limited to the particular shapes of regions illustrated herein but are to include deviations in shapes
25 that result, for example, from manufacturing. For example, an implanted region illustrated as a rectangle will, typically, have rounded or curved features and/or a gradient of implant concentration at its edges rather than a binary change from implanted to non-implanted region. Likewise, a buried region formed by implantation may result in some implantation in the region between the buried region and the
30 surface through which the implantation takes place. Thus, the regions illustrated in the figures are schematic in nature and their shapes are not intended to illustrate the precise shape of a region of a device and are not intended to limit the scope of the present invention.

Embodiments of the present invention provide ohmic contacts on re-grown contact regions of a Group III-nitride based transistor and methods of forming such contacts.

Embodiments of the present invention may be suited for use in nitride-based HEMTs such as Group III-nitride based devices. As used herein, the term "Group III nitride" refers to those semiconducting compounds formed between nitrogen and the elements in Group III of the periodic table, usually aluminum (Al), gallium (Ga), and/or indium (In). The term also refers to ternary and quaternary compounds such as AlGaN and AlInGaN. As is well understood by those in this art, the Group III elements can combine with nitrogen to form binary (*e.g.*, GaN), ternary (*e.g.*, AlGaN, AlInN), and quaternary (*e.g.*, AlInGaN) compounds. These compounds all have empirical formulas in which one mole of nitrogen is combined with a total of one mole of the Group III elements. Accordingly, formulas such as $\text{Al}_x\text{Ga}_{1-x}\text{N}$ where $0 \leq x \leq 1$ are often used to describe them.

Suitable structures for GaN-based HEMTs that may utilize embodiments of the present invention are described, for example, in commonly assigned U.S. Patent 6,316,793 and U.S. application serial no. 09/904,333 filed July 12, 2001 for "ALUMINUM GALLIUM NITRIDE/GALLIUM NITRIDE HIGH ELECTRON MOBILITY TRANSISTORS HAVING A GATE CONTACT ON A GALLIUM NITRIDE BASED CAP SEGMENT AND METHODS OF FABRICATING SAME," U.S. provisional application serial no. 60/290,195 filed May 11, 2001 for "GROUP III NITRIDE BASED HIGH ELECTRON MOBILITY TRANSISTOR (HEMT) WITH BARRIER/SPACER LAYER" and United States Patent Application Serial No. 10/102,272, to Smorchkova *et al.*, entitled "GROUP-III NITRIDE BASED HIGH ELECTRON MOBILITY TRANSISTOR (HEMT) WITH BARRIER/SPACER LAYER" the disclosures of which are hereby incorporated herein by reference in their entirety.

Fabrication of embodiments of the present invention is schematically illustrated in **Figures 1A-1G**. As seen in **Figure 1A**, a substrate **10** is provided on which nitride based devices may be formed. In particular embodiments of the present invention, the substrate **10** may be a semi-insulating silicon carbide (SiC) substrate that may be, for example, 4H polytype of silicon carbide. Other silicon carbide candidate polytypes include the 3C, 6H, and 15R polytypes. The term "semi-

insulating" is used descriptively rather than in an absolute sense. In particular embodiments of the present invention, the silicon carbide bulk crystal has a resistivity equal to or higher than about $1 \times 10^5 \Omega\text{-cm}$ at room temperature.

Optional buffer, nucleation and/or transition layers (not shown) may be provided on the substrate 10. For example, an AlN buffer layer may be provided to provide an appropriate crystal structure transition between the silicon carbide substrate and the remainder of the device. Additionally, strain balancing transition layer(s) may also be provided as described, for example, in commonly assigned United States Patent Application Serial No. 10/199,786, filed July 19, 2002 and entitled "STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTORS AND METHODS OF FABRICATING STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTORS, and United States Provisional Patent Application Serial No. 60/337,687, filed December 3, 2001 and entitled "STRAIN BALANCED NITRIDE HETEROJUNCTION TRANSISTOR," the disclosures of which are incorporated herein by reference as if set forth fully herein.

Silicon carbide has a much closer crystal lattice match to Group III nitrides than does sapphire (Al_2O_3), which is a very common substrate material for Group III nitride devices. The closer lattice match may result in Group III nitride films of higher quality than those generally available on sapphire. Silicon carbide also has a very high thermal conductivity so that the total output power of Group III nitride devices on silicon carbide is, typically, not as limited by thermal dissipation of the substrate as in the case of the same devices formed on sapphire. Also, the availability of semi-insulating silicon carbide substrates may provide for device isolation and reduced parasitic capacitance. Appropriate SiC substrates are manufactured by, for example, Cree, Inc., of Durham, N.C., the assignee of the present invention, and methods for producing are described, for example, in U. S. Patent Nos. Re. 34,861; 4,946,547; 5,200,022; and 6,218,680, the contents of which are incorporated herein by reference in their entirety. Similarly, techniques for epitaxial growth of Group III nitrides have been described in, for example, U. S. Patent Nos. 5,210,051; 5,393,993; 5,523,589; and 5,292,501, the contents of which are also incorporated herein by reference in their entirety.

Although silicon carbide may be the preferred substrate material, embodiments of the present invention may utilize any suitable substrate, such as sapphire, aluminum nitride, aluminum gallium nitride, gallium nitride, silicon, GaAs,

LGO, ZnO, LAO, InP and the like. In some embodiments, an appropriate buffer layer also may be formed.

Returning to **Figure 1A**, a channel layer **20** is provided on the substrate **10**. The channel layer **20** may be deposited on the substrate **10** using buffer layers,
5 transition layers, and/or nucleation layers as described above. The channel layer **20** may be under compressive strain. Furthermore, the channel layer and/or buffer nucleation and/or transition layers may be deposited by MOCVD or by other techniques known to those of skill in the art, such as MBE or HVPE.

In some embodiments of the present invention, the channel layer **20** is a Group
10 III-nitride, such as $\text{Al}_x\text{Ga}_{1-x}\text{N}$ where $0 \leq x < 1$, provided that the bandgap of the channel layer **20** is less than the bandgap of the barrier layer **22**. In certain embodiments of the present invention, $x = 0$, indicating that the channel layer **20** is GaN. The channel layer **20** may also be other Group III-nitrides such as InGaN, AlInGaN or the like. The channel layer **20** may be undoped ("unintentionally doped")
15 and may be grown to a thickness of greater than about 20 Å. The channel layer **20** may also be a multi-layer structure, such as a superlattice or combinations of GaN, AlGaN or the like.

A barrier layer **22** is provided on the channel layer **20**. The channel layer **20** may have a bandgap that is less than the bandgap of the barrier layer **22**. The barrier
20 layer **22** may be deposited on the channel layer **20**. In certain embodiments of the present invention, the barrier layer **22** is AlN, AlInN, AlGaN or AlInGaN with a thickness of between about 1 and about 100 nm. In some embodiments of the present invention, the barrier layer **22** includes multiple layers. For example, the barrier layer **22** may be about 1 nm of AlN with about 25 nm of AlGaN on the AlN layer.
25 Examples of barrier layers according to certain embodiments of the present invention are described in United States Patent Application Serial No. 10/102,272, to Smorchkova *et al.*, entitled "GROUP-III NITRIDE BASED HIGH ELECTRON MOBILITY TRANSISTOR (HEMT) WITH BARRIER/SPACER LAYER" the disclosure of which is incorporated herein by reference as if set forth fully herein.

30 The barrier layer **22** may be a Group III-nitride and has a bandgap larger than that of the channel layer **20**. Accordingly, in certain embodiments of the present invention, the barrier layer **22** is AlGaN, AlInGaN and/or AlN or combinations of layers thereof. Other materials may also be used for the barrier layer **22**. For

example, ZnGeN₂, ZnSiN₂ and/or MgGeN₂ could also be used. The barrier layer 22 may, for example, be from about 1 to about 100 nm thick, but is not so thick as to cause cracking or substantial defect formation therein. Preferably, the barrier layer 22 is undoped or doped with an n-type dopant to a concentration less than about 10¹⁹ cm⁻³. In some embodiments of the present invention, the barrier layer 22 is Al_xGa_{1-x}N where 0 < x < 1. In such embodiments, the barrier layer 22 may be from about 3 to about 30 nm thick. In particular embodiments, the aluminum concentration is about 25%. However, in other embodiments of the present invention, the barrier layer 22 comprises AlGaN with an aluminum concentration of between about 5% and about 100%. In specific embodiments of the present invention, the aluminum concentration is greater than about 10%. In embodiments of the present invention where the barrier layer 22 comprises an AlN layer, the thickness of the barrier layer 22 may, for example, be from about 0.3 nm to about 4 nm.

Figure 1B illustrates formation of an optional first dielectric layer 24. The first dielectric layer 24 may be a silicon nitride layer, such as an Si_xN_y layer. The silicon nitride layer may serve as a passivation layer for the device. The silicon nitride layer may be deposited by, for example, plasma-enhanced chemical vapor deposition (PECVD), low pressure chemical vapor deposition (LPCVD) and/or sputtering. The silicon nitride layer may be deposited in the same reactor as other layers of the transistor. In some embodiments of the present invention, other dielectrics may also be utilized, such as, for example, silicon oxynitride and/or silicon dioxide.

Figure 1C illustrates formation of a mask 30 on the first dielectric layer 24. The mask 30 is formed on the region of the barrier layer 22 where a gate contact will subsequently be formed. As illustrated in **Figure 1C**, the wafer of **Figure 1C** may be removed from the epi reactor and patterned with a mask material 30 to expose the desired recess areas. The mask material 30 should be able to withstand the growth temperature of subsequent processing, including the formation of regrown contact regions 26 as described below. In particular embodiments of the present invention, the mask 30 is provided by an oxide. In certain embodiments of the present invention, the mask 30 is patterned using lift-off techniques. Alternatively, a wet or dry etch could be utilized to pattern the mask 30. In certain embodiments of the present invention, SiO_x is the mask material, although other materials, such as AlN and Si_xN_y based materials, may also be used. A photo-resist, e-beam resist material or

organic mask material may also be utilized if it is not unduly damaged by subsequent processing steps, such as deposition temperatures or the like.

As illustrated in **Figure 1D**, after formation and patterning of the mask **30** so as to leave the mask material in the regions where contact recesses **23** are not to be formed for the ohmic contacts, the recesses are etched through the first dielectric layer **24**, through the barrier layer **22**, to the channel layer **20** and, in some embodiments, to and into the channel layer **20** or, in some embodiments, even through the channel layer **20**. The etch to form the contact recesses **23** may be provided by, for example, a wet etch, a dry etch and/or a reactive ion etch or the like. Optionally, the structure may be annealed to remove and/or reduce damage resulting from the etch. Furthermore, optionally, the periphery of the device may be etched to form a mesa structure (not shown), for example, if other termination structures, such as a field plate, implant or other termination structure are not provided or may be provided in addition to such structures.

As illustrated in **Figure 1E**, after etching the contact recesses **23**, a contact layer **26'** is formed on the exposed regions of the channel layer **20** and the mask **30**. For example, the wafer of **Figure 1D** may be put back into the epi reactor for deposition of the contact layer **26'**. In some embodiments of the present invention, the contact layer **26'** is formed by metal organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), plasma enhanced chemical vapor deposition (PECVD), sputtering and/or hydride vapor phase epitaxy (HVPE). In some embodiments, the contact layer **26'** is regrown at a reduced deposition temperature. In particular, a low temperature deposition process may be used. As used herein "low temperature deposition" refers to formation of a layer at a temperature lower than a temperature at which substantial mass transport from the wafer to the regrown region takes place. For example, the contact layer **26'** may be formed at a temperature of from about room temperature to about 950 °C. In some embodiments, the contact layer **26'** is formed at a temperature of less than 960 °C. In particular embodiments of the present invention, the contact layer **26'** is formed at a very low temperature, for example, at a temperature of less than about 450 °C and in some embodiments, at temperature of less than about 200 °C. Such very low temperature conditions may be used, for example, with sputtering and/or PECVD growth techniques. The use of a reduced deposition temperature and/or low temperature deposition may reduce trapping and/or may provide improved reliability.

As discussed below, the contact layer 26' may be unevenly formed such that the portion on the mask 30 is porous or discontinuous. In some embodiments, the contact layer 26' is not formed on the mask 30. Such a selective formation of the contact layer 26' may depend on the composition of the contact layer 26', the mask 30 and the growth conditions for the contact layer 26'.

In some embodiments of the present invention, the contact layer 26' may be an n-type degenerate semiconductor material. In certain embodiments of the present invention, the contact layer 26' may be heavily doped n-type InN, InAlN, AlGaIn, AlInGaIn, GaN and/or InGaIn. In other embodiments of the present invention, the contact layer 26' may be an n-type degenerate semiconductor material other than GaN or AlGaIn. For example, the contact layer may be a non-nitride Group III-V semiconductor material, a Group IV semiconductor material and/or a Group II-VI semiconductor material. Examples of potential contact layer 26' materials include, for example, ZnO, ZnGeN₂ and/or ZnSnN₂. In other embodiments of the present invention, the contact layer 26' may be a metal or metal alloy, for example, a metal silicide, capable of conformal deposition at a low temperature that has a low work function and does not form a Schottky contact. For example, an MOCVD conformal deposition of Al using DMAIH at a temperature of at least about 200 °C to touch the 2DEG. The metal may be subsequently etched off in the channel and gate regions. Furthermore, a passivation layer may be deposited before deposition of the metal.

Formation of a GaN contact layer 26' may reduce and/or eliminate a band discontinuity with the channel layer 22 if the channel layer 22 is also GaN. The contact layer 26' is formed to a thickness sufficient to provide a low sheet resistivity. For example, the contact layer 26' may be grown to a thickness sufficient to provide a sheet resistivity that is less than a sheet resistivity of the 2DEG formed at the interface between the channel layer 20 and the barrier layer 22. Several tens of nanometers of GaN, for example, may be sufficient thickness for the contact layer 26', however, a thicker layer may have a lower resistances and increase transfer length (L_T). The contact layer 26' may be doped with Si, Ge and/or O or other suitable n-type dopant or may be naturally n-type as deposited. The contact layer 26' may be doped as formed rather than through subsequent ion implantation. Formation of the doped contact layer 26' without ion implantation may avoid the need for extremely high temperature annealing to activate the dopants. In particular embodiments of the present invention, the contact layer 26' has a sheet resistivity of from about 10 to

about 400 Ω/\square . In certain embodiments of the present invention, the contact layer is doped to provide a carrier concentration of from about 10^{18} to about 10^{21} cm^{-3} . In still further embodiments of the present invention, the contact layer 26' is from about 10 nm to about 1000 nm thick.

5 For a contact layer 26' that is not a Group III nitride material, the mask 30 may not be needed as the material could be blanket deposited and then patterned and etched after deposition.

Figure 1F illustrates removal of the portion of the contact layer 26' formed on the mask 30 and removal of the mask 30 to expose the first dielectric layer 24, thus
10 providing the contact regions 26. The mask 30 and the portion of the contact layer 26' may be removed, for example, by etching the mask 30 in buffered HF or other etchant that will remove the mask layer 30 and leave the first dielectric layer 24 and the contact regions 26. In some embodiments of the present invention, the portion of the contact layer 26' is formed on the mask 30 may be formed so as to be porous or
15 discontinuous so as to allow etching the mask 30 through the contact layer 26'. In such a way, the mask 30 may be etched with an etchant that has etching selectivity with respect to the contact layer 26' and the first dielectric layer 24. For smaller geometry devices, the mask layer 30 may be etched from the sides to remove the mask layer and the portion of the contact layer 26' on the mask layer 30, for example,
20 if the portion of the contact layer 26' on the mask layer is not porous or is continuous.

Figure 1G illustrates formation of a gate recess in the first dielectric layer 24 and formation of a gate contact 44 on the exposed portion of the barrier layer 22 in the gate recess. For example, a gate recess may be etched through the first dielectric layer 24 using, for example, a dry etch, a wet etch and/or RIE or the like. Optionally,
25 the structure may be annealed to repair some or all of the damage resulting from the etch of the gate recess. Suitable gate contact materials include, for example, Ni, Pt, Pd or other such Schottky contact materials. Additional overlayers may also be provided. In some embodiments of the present invention, the gate contact 44 may be formed on the dielectric layer 24.

30 As is further seen in Figure 1G, ohmic contacts 40 and 42 are formed on the contact regions 26 and may provide source and drain contacts. The ohmic contacts 40 and 42 may be formed before or after formation of the gate recess and/or contact 44. In some embodiments of the present invention, the ohmic contacts 40 and 42 are annealed, for example at a temperature of about 850 °C or less. In other

embodiments, the anneal of the ohmic contacts is not carried out. The use of a reduced anneal temperature or no anneal may reduce trapping and/or may provide improved reliability. The presence of the highly doped n-type contact regions may lower contact resistance that may provide for increased efficiency and/or radio frequency power density. Suitable ohmic contact materials include, for example, a Ti/Al/Ni/Au stack may be used. Similarly, a structure of Ti/Al/X/Au may be used where X may be Mo, Pt and/or Ti.

While embodiments of the present invention have been described with reference to a blanket deposition of the contact layer 26', alternatively, selective regrowth of the contact regions 26 could also be utilized while still benefiting from the teachings of the present invention. Furthermore, the regrown contact regions 26 may be provided for only one of the ohmic contacts 40 and 42 and a conventional contact structure provided for the other contact. Accordingly, embodiments of the present invention should not be construed as limited to the specific processing steps illustrated in **Figures 1A-1G**.

Furthermore, while a specific sequence of processing steps has been described, deviation from this sequence may be provided while still being within the scope of embodiments of the present invention. For example, the gate recess and gate contact may be formed before or after formation of the ohmic contacts or even before formation of the contact recesses. Accordingly, embodiments of the present invention should not be construed as limited to the specific sequence of operations described above.

Figures 2, 3A and 3B are illustrations of embodiments of the present invention incorporating contact area enlargement structures that provide increased vertical surface area of the interface between the contact region 26 and the channel layer 20. **Figure 2** illustrates embodiments of the present invention as incorporating a contact area enlargement area structure in a sidewall of a portion of the channel layer 20 and **Figures 3A and 3B** illustrate embodiments of the present invention where the contact enlargement area structure is provided by filled holes extending into the contact layer 20. While each of the contact area enlargement structures is described herein separately, the contact area enlargement structures may also be provided in combination with each other or other structures that increase the vertical contact area between the channel layer 20 and the contact region 26 as compared to a planar vertical contact area. Such structures may provide means for increasing a surface area

of an interface between a vertical portion of the n-type nitride-based semiconductor material contact region 26 and the nitride-based channel layer 20.

Figures 2, 3A and 3B illustrate a partial section of a transistor illustrating a single ohmic contact region. As will be appreciated by those of skill in the art, a corresponding section may be provided for a second ohmic contact region opposite the gate contact so as to provide source and drain contacts. Alternatively, embodiments of the present invention may provide a contact area enlargement structure for only one of the ohmic contacts.

Figure 2 is a top view of a portion of a HEMT according to further embodiments of the present invention. As seen in **Figure 2**, the surface area of the interface between the contact region 26 and the channel layer 20 and/or the barrier layer 22 may be increased by providing an increased surface area sidewall 200 of the channel layer 20 and/or the barrier layer 22. The increased surface area sidewall 200 has an increased surface area with respect to a straight sidewall. Increasing the surface area of the interface between the contact region 26 and the channel layer 20 may reduce the resistance between the contact region 26 and the channel layer 20.

The patterned sidewall 200 may be provided by patterning the channel layer 20 during the contact recess etch described above. For example, a first etch through the first dielectric layer 24 and the barrier layer 22 may be performed and then a second etch into the channel layer 22 may be performed with a mask on the exposed portion of the channel layer 20 to provide the pattern of the sidewall. Alternatively, a single etch may be performed if the mask 30 has a pattern corresponding to the desired sidewall pattern of the channel layer 20.

The sidewall may have a regular or irregular repeating or non-repeating shape. The sawtooth shape illustrated in **Figure 2** is provided as an example of a shape that may be used. However, other shapes may also be used, for example, a notch shape, a series of curves or the like may be used. Accordingly, some embodiments of the present invention should not be limited to a particular shape for the increased surface area sidewall 200.

The ohmic contact metal 42' is also illustrated in **Figure 2** on the contact region 26. The contact metal 42' is illustrated as stopping before the periphery of the contact region 26. However, the contact metal 42' may extend further than illustrated and may, for example, extend onto the channel layer 22.

Figure 3A is a top view and **Figure 3B** is a cross-section taken along the lines I-I' of **Figure 3A** of further embodiments of the present invention. As seen in **Figure 3**, the surface area of the interface between the regrown contact region **26** and the channel layer **20** may be increased by providing holes **300** that extend into the channel layer **20**. The holes **300** have n-type material in them as provided in the contact region **26**. The ohmic contact **42''** extends to cover the holes **300** so that the n-type material in the holes **300** is electrically connected to the contact region **26**.

The filled holes **300** may be provided by patterning the channel layer **20** during the contact recess etch described above so as to provide holes that are present when the contact layer **26'** is formed. For example, a first etch through the first dielectric layer **24** and the barrier layer **22** may be performed and then a second etch into the channel layer **22** may be performed with a mask on the exposed portion to the channel layer **20** to provide the holes.

Alternatively, a single etch may be performed if the mask **30** has a pattern corresponding to the desired holes of the channel layer **20**. In such a case, the holes would extend through the barrier layer **22** and to or into the channel layer **20**. The contact metal would then extend onto the barrier layer **22** to contact the material in the holes **300** as illustrated in **Figure 3B**.

The holes **300** may have a regular or irregular repeating or non-repeating pattern. Furthermore, the holes **300** may also have a circular or other shape periphery. The pattern of holes and shape of holes illustrated in **Figure 3A** is provided as an example of a pattern and shape that may be used. However, other patterns and/or shapes may also be used. Accordingly, some embodiments of the present invention should not be limited to a particular pattern and/or shape for the holes **300**.

Figures 4A-4C illustrate fabrication of further embodiments of the present invention where a contact region is provided that extends onto the barrier layer. Fabrication of the embodiments of the present invention illustrated in **Figures 4A-4C** may be the same as that illustrated in **Figures 1A-1F** except that the first dielectric layer **24** is resized to a smaller size first dielectric layer **424** so as to expose a portion of the barrier layer **22**, for example, by undercutting the mask **30** into the first dielectric layer **24** with an isotropic etch. Alternatively, the mask **30** could be stripped and another mask applied and the first dielectric layer **24** may be etched using this second mask. While the first dielectric layer **424** in **Figures 4A** is described herein as

a dielectric material, other removable materials that may withstand the conditions for deposition of the contact regions may be used.

As seen in **Figure 4A**, the contact region **426** is regrown as discussed above and the mask **30** is removed. As seen in **Figure 4B**, the first dielectric layer **424** is removed and a second dielectric layer **430** is conformally deposited on the contact layers **426** and the barrier layer **22**. The second dielectric layer **430** would typically be deposited isotropically. Windows in the second dielectric layer **430** may be provided on the contact layer **426** and ohmic contacts for the source and drain contacts **440** and **442** may be formed on the contact layer **426**. The ohmic contacts may also be formed prior to the deposition of the second dielectric layer **430**.

As seen in **Figure 4C**, the second dielectric layer **430** is anisotropically etched to expose the barrier layer **22** and provide sidewall spaces **430'** and a gate contact recess. A gate metal may be deposited and patterned, for example, using lift-off techniques, to provide the gate contact **444**. The length of the gate contact **444** may be approximately the width of the first dielectric layer **424** less twice the thickness of the second dielectric layer **430**. In particular embodiments of the present invention, the first dielectric **424** may have a width of about 0.5 to about 1 μm and the second dielectric **430** may have a thickness of from about 0.1 to about 0.5 μm .

While embodiments of the present invention have been illustrated with regrown contact regions for both the source and drain contacts, such regrown regions may be provided for only one of the source or the drain. Furthermore, while the gate contacts have been illustrated as substantially centered between the source and drain, in certain embodiments of the present invention, the gate contact may be offset, for example, toward the source contact.

Furthermore, while embodiments of the present invention have been described with reference to a particular sequence of fabrication steps, a different sequence of steps may be utilized while still falling within the scope of the present invention. Accordingly, embodiments of the present invention should not be construed as limited to the particular sequence of steps described herein.

In the drawings and specification, there have been disclosed typical embodiments of the invention, and, although specific terms have been employed, they have been used in a generic and descriptive sense only and not for purposes of limitation, the scope of the invention being set forth in the following claims.